AFRL-OSR-VA-TR-2013-0003

Co-evolution of Knowledge and Event Memory

Richard M Shiffrin

Indiana University

FEBRUARY 2013
Final Report

DISTRIBUTION A: Approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSL
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
19-02-2013	Final Performance Report	From 01-03-2009 to 30-11-2011

4. TITLE AND SUBTITLE Co-evolution of Knowledge and Event Memory	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER FA9550-09-1-0178
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Richard M Shiffrin	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Trustees of Indiana University 1101 E. Tenth St. Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405-7007	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/RSL 801 N. Randolph St. Arlington, VA 22203-1977 Dr. Jay Myung	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-OSR-VA-TR-2013-0003

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION A: APPROVED FOR PUBLIC RELEASE

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The project investigated the inter-relation of event memory and knowledge as they co-evolve together. The formation of knowledge is based on an accumulation of information from many similar events. The coding of events is based on the then-current status of related knowledge. Key empirical studies established the role of frequency of events in storage and retrieval, the role of experimental and pre-experimental memory traces, statistical accumulation of information across many events, and the storage of new traces for test events. The key element of the theory, abbreviated SARKAE (for the Storage And Retrieval of Knowledge And Events), is the assumption that that an event occurrence produces two sorts of memory storage: an incomplete and error prone event trace, and addition of some of this same information to an existing trace that is sufficiently similar (including earlier event traces, and any developed knowledge trace). The latter addition is the process that allows knowledge to grow from events.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Richard M Shiffrin
					19b. TELEPHONE NUMBER *(Include area code)* 812-855-4972

Reset

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Performance Report

AFOSR Grant FA9550-09-1-0178
for period 3/1/2009-11/30/2011

Title: Co-evolution of Knowledge and Event Memory

Program manager: Dr. Jun Zhang

PI: Richard M Shiffrin
Institution: Indiana University Bloomington
Date of submission: February 19, 2013

0. **Preamble and Introduction**

The grant title was the "Co-Evolution of Knowledge and Event Memory". The main research effort was to carry out empirical research and develop a model that would explicate the way that lifetime events produce our general knowledge and the way that general knowledge is used to code events as they occur. Although it is evident that there must be a close linkage between these two forms of memory, and essentially all our performance in any task is determined by the processes that link these two, the subject has not heretofore been the subject of much research and modeling.

It is important not to confuse this issue with the dualism between short-term and long-term memory, which is also an important component of the developed theory. An event enters short-term memory via an interaction with knowledge, and resides there for a time, as coding and storage takes place. Two primary types of storage occur: Storage of an incomplete and error prone copy of the information in short-term memory (i.e. storage of the event in long-term memory), and additional storage in a previously stored long-term memory trace that is sufficiently similar. When the previous trace is, say, the first occurrence of the same information, then the additional storage becomes the first step in the formation of what eventually will be a knowledge trace. E.g. the first occurrence of a word causes storage of an event trace; the next occurrence causes storage of another event trace but also additional storage in the previous trace, thus becoming an early step in the eventual formation of the lexical trace for that word. Eventually a knowledge trace becomes replete with a great deal of information, centered around the information that remains consistent across many storage events. Thus a lexical trace for a word comes to center around the common meaning, spelling and phonology across events, but the context that is stored in event trace, and that accumulates in the lexical trace, keeps changing and hence acts more like noise than common content. This idea explains how knowledge comes to appear context free, not associated with a particular lifetime event.

Of course knowledge traces (whether words, faces, game playing, or motor actions) develop relatively slowly across long periods of time and many event occurrences. Thus there is a continuum of traces from 'weak' event traces at one end to 'rich' knowledge traces at the other. At any stage of this development the then existing knowledge traces (and event traces to some degree as well) are used to code new events--every event must be coded in terms of what we already have learned. This is the reason for the term 'co-evolution'.

We explored these ideas in a number of empirical studies, including ones that investigated the critical role of event frequency using novel stimuli so that the findings would be minimally contaminated by prior learning. The single main result of the research is a long empirical and theoretical article that was developed during the course of the grant, with the same title as the grant title, that is just now in press in Psychological Review. A number of other studies were carried out to explore other aspects of the theory, and these will be described in the body of this report.

1. **Scientific Objectives of Research**

Much of the memory research over the past hundred years has been focused on the storage and retrieval of events. Most often events have been represented as words presented on a study list, although some research has focused on other materials such as faces, scenes, or auditory stimuli. The PI has been a major contributor to this growing body of results and theory (e.g. Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1980, 1981; Shiffrin & Steyvers, 1997, among many other publications). This research has produced the 'standard' models of event recall and event recognition. In most of the research on event memory it is recognized either implicitly or sometimes explicitly that retrieval from general knowledge is important generally, and also for the purpose of encoding events. Thus retrieval of developed general knowledge has been incorporated in some models, but the way that such general knowledge forms has not been explored as fully. (We note that some research has occurred in the form of neural net models, but the aim has been quite different, mainly focused on developed word knowledge, and high level sematic information, and not aimed at memory per se. This report will not discuss such research further, but we note that the relation of the approaches is discussed in Nelson & Shiffrin, in press).

There has been some important research relevant to the interaction between event memory and knowledge, in the form of long-term priming: Occurrence of an event changes long term retrieval of knowledge about the contents of that event. For example, study of a word on a list speeds later long term retrieval from the lexical trace for that word, measured by naming or lexical decision, and improves accuracy of identification of a briefly flashed presentation of that word. Such effects occur well outside the range of short term memory, typically at test delays of 20 minutes or more. Shiffrin and Steyvers (1997, 1998) proposed a mechanism for such priming: When a word is studied on a list, it causes not only storage of an event trace but additional storage in that word's knowledge trace. That additional storage includes current context, and when a subsequent test (say, naming) occurs in a similar context (e.g. the laboratory) the match of the current context cue to the augmented knowledge trace is higher and produces faster and better knowledge trace retrieval. This idea forms the starting basis for the research in the AFOSR grant.

Even less research has been devoted to the way that events are coded in terms of then existing general knowledge. It is evident that the features we use to code events are retrieved from general knowledge in long-term memory, but how development of that general knowledge affects such coding is largely unknown. Thus the present research project aims to fill these gaps in our understanding of memory process, under the general rubric of co-evolution of knowledge and event memory.

In particular, the empirical research explores what be the most important component of such co-evolution, the frequency with which similar events re-occur, and the similarity of the context in which re-occurrences take place. We therefore trained novel events and knowledge over several weeks: Some learners studied Chinese characters (with which they were unfamiliar) in a visual search task while others studied these stimuli in a character matching task. During training, stimulus frequency and contextual variability was varied substantially across the Chinese characters. All characters (and new ones) were subsequently tested in three tasks. One was an event memory task, known as episodic recognition: A list of Chinese characters was studied and then tested for old/new recognition. Two tests of retrieval from general knowledge were employed: 1) Pseudo-lexical decision in which trained and untrained Chinese characters were presented and the learners tried to classify these; 2) Forced-choice perceptual identification in which a briefly flashed and masked character was followed by two choices one of which had been flashed. The purpose of such studies was in part to ascertain the basis for the always observed frequency effects in all memory tasks (e.g. worse performance for high frequency words in recognition and better performance for high frequency words in knowledge retrieval). We wanted to know in particular whether such effects were due directly to the extra occurrences of the corresponding events, or instead due to the differing contexts in which high and low frequency events tend to occur (e.g. a high frequency word like 'song' may occur in many different life settings, whereas a low frequency word like 'scalpel' may always occur in a medical/surgical context). Thus we used two training tasks: The visual search task had high frequency characters co-occurring in training with other high frequency characters, whereas the character matching tasks had characters of any frequency only occurring with themselves, because two successive presentations of the same character had to be judged for slight variations in contrast, orientation, or size.

A number of other investigations explored other aspects of event memory and knowledge, as described in section 3 below.

2. **Technical Approach**

The empirical studies are designed with careful control of stimuli and training (novel stimuli so that number of events experienced can be controlled and learning assessed). Even more important the number of tasks used covers a wide span. Almost all studies and models for the results are embedded in variants of a single paradigm, but the way that the developed models can apply across paradigms is left as a promissory note. In our research, five quite different tasks are used, tasks that span the kinds of paradigms used in learning and memory research: Visual search, character matching, episodic recognition memory, 'lexical' decision, and perceptual identification. A key component of the research effort is the development of a quantitative model in which the key components apply to all these very different tasks.

The basic research we have outlined explores how knowledge develops from events, and hence the learning is basically statistical in nature: One event has many possible relations among the information that is encoded and present in short term memory. It is through the accumulation of many events that the critical information structure emerges. Thus a second line of research explored statistical learning, using paradigms in which associations are to be learned but the information present on any one trial is ambiguous. The research investigated cross modal associative learning: A typical trial contains four novel pseudo-words presented successively and auditorily, and four novel object pictures. Which objects go with which sounds is not provided and thus is ambiguous. However, as these and various other sounds and pictures accumulate on successive trials, the correct pairings emerge statistically. This line of research adds important knowledge and constraints for the developing co-evolution theory. We also produced a quantitative model to explain the main processes at work in such statistical learning.

Another key element of the theory is the necessity of (multiple) storage of each event, including events that occur at each test trial. This motivated another series of studies investigating the effects of storage of test trials. It is well known that repeated testing of a given item is beneficial, especially at appropriate delays, but we looked at the effects of testing upon other items yet to be tested. Such storage of test events should harm subsequent retrieval due to competition. We showed such effects in several studies, and modeled the results.

3. **Progress Made and Results Obtained**

The central core of the research accomplished in the grant period is encapsulated in a long article now in press in Psychological Review with the title The Co-evolution of Knowledge and Event Memory (co-authored with my then graduate student, Angela Nelson; this article was first submitted at about the time the grant ended, but the process of revising twice delayed publication). This brief summary can only hint at the findings and theory contained in the 40 pages in the journal, let alone the additional research completed during the grant period.

Empirical Findings:

a) Training: Visual search training had Os look for the presence or absence of a Chinese character target in a display of two or four such characters. Varied training was used so that targets on some trials were foils on another, and vice versa. Character matching (Experiment 2) required Os to decide whether a Chinese character presented twice in succession was identical or changed slightly in orientation, size, or contrast. In each training paradigm, two weeks of daily sessions produced learning that was greater for characters presented more often, measured both by task performance increases, and subsequent transfer performance.

b) **Transfer**:

1) Episodic Recognition (Figures?): When testing followed immediately after list study, in both experiments, higher frequency characters were recognized less well, and the results showed a mirror effect, with more difficult conditions having lower hit rates and higher false alarm rates (matching prior findings with words). In Experiment 2 (with character matching training) recognition study and test was also carried out after a one-month delay. The delayed results showed greatly reduced (possibly missing) effects of training frequency on recognition. The Experiment 2 findings occurred despite training that eliminated similarity differences between characters based on differential training. Such findings had important implications for modeling recognition, as detailed below.

2) Forced choice perceptual identification (Experiment 1 only): A character is briefly presented and masked, followed by a choice between the presented item (the target) and a foil. The higher the frequency of the target the higher the probability of correct choice, and the higher the frequency of the foil, the higher the probability of correct choice. The former result matches prior findings for words, but in prior word research, frequency of foils tends to have little effect. (Our results are perhaps easier to interpret given that so many uncontrolled factors co-vary with word frequency). To explain our foil findings we relied on a discounting explanation, as described below.

3) Pseudo-lexical decision: A trained or untrained character was presented on each trial for a decision concerning whether it had been (ever) trained. Very strong effects of training frequency (faster

responses for higher frequency) were observed in both experiments. In addition, in Experiment 2, the frequency effects were equally pronounced with immediate or delayed testing.

Theory and Modeling: The general theory was termed SARKAE, for Storage and Retrieval of Knowledge and Events. It built on prior theories (by the PI) known as SAM and REM, but represented a substantial generalization, and opened much new ground. A very simplified version of SARKAE was used to produce a quantitative simulation of the above findings, with the goal of exhibiting the common processes at work across the various tasks, and showing how even a simplified theory with those common elements could produce the qualitative trends that were observed. Various more realistic and complex extensions of the theory and potential applications to other results and tasks were taken up in a long general discussion.

1) Common processes:

Representation: All memory probes and traces in memory are represented as vectors of feature classes (e.g. 'color') that are each subdivided into binary feature values (e.g. presence or absence of 'red'). The features and values include general task and environmental context, features from events and knowledge residing in short-term memory in close temporal proximity, and features directly representing the present event (e.g. physical and semantic features representing 'chair' if 'chair' is presented). For simplicity event traces are assumed to have at most one feature value count per feature. Knowledge traces grow in 'richness' as similar events re-occur, and come to have multiple counts per feature value, multiple values per feature, and increasing numbers of features with counts. These features are stored in incomplete and error prone fashion in event traces, and in error prone but accumulating fashion in knowledge traces. The features extracted (retrieved) from both knowledge traces and event traces are entered into a continuously changing short-term memory representation..

Coding: A test probe arrives in the form of sensory information, and these are added to pre-existing information such as environmental and internal context, including features left over from recent events. These features are used to probe knowledge. Retrieval from knowledge produces features that are added (over time) to those already in short-term memory. This process is incomplete (only some of the many directly relevant features in knowledge are extracted) and error prone (occasionally a feature is retrieved on the basis of general base rates in knowledge rather than from the relevant knowledge trace). If the task requires a knowledge decision, it is these features and the dynamic changes in those features that are used to make a decision. If the task requires retrieval from event traces, then these features are used as a probe that causes incomplete and error prone retrieval (evolving over time) from event traces. The features retrieved from event traces are used to make episodic decisions (often augmented by 'sophisticated guessing' based on general knowledge).

Retrieval and Frequency: Knowledge traces vary enormously in 'richness', defined as the accumulation of counts of feature values in multiple features. Richness does affect retrieval, in two primary ways. One has been defined as 'context variability', which is a bit misleading in present terms because it refers not so much to general task and situational context as the variability of other events that co-occur with a given event or knowledge trace. Thus a high frequency word may occur in many different kinds of settings, thereby co-occurring with many other high frequency words, but a low frequency word may occur mainly in a single setting or scenario. Because the features in short-term memory that encode an event include features from other items/events in short-term memory at the same time, and these features form the basis for storage in event traces and knowledge traces, the theory incorporates a critical role for 'context variability': Retrieval of high frequency items incorporates retrieval of features of co-occurring items. This

factor operates for example in Experiment 1, because the visual search task used training in which high frequency characters co-occurred with other high frequency characters. Is this the only factor that produces frequency effects? We tested this in Experiment 2, in which training used character matching so that the character-context for any character was only itself. Nonetheless frequency effects were found and in some cases were quite strong (e.g. pseudo-lexical decision). Thus a role for 'pure frequency' had to be incorporated in the model: It was assumed that retrieval from a knowledge trace was less error prone to the degree that the trace was 'richer'. Thus the model that was fit to the data from the various studies included both factors that are based on frequency: 'context variability' and 'trace richness'.
Retrieval as a Bayesian Likelihood Ratio: A set of various kinds of features is used to probe memory. For both event and knowledge traces, it is assumed that all traces in memory that are sufficiently rich and sufficiently similar to the probe are compared in parallel. Each comparison of probe to trace is based on feature and feature value matches and mismatches, and produces a likelihood ratio that is the SARKAE equivalent of strength or activation value. The likelihood ratio gives the probability that probe and trace were produced by the same event divided by the probability that probe and trace were produced by different events. These likelihood ratios are used to govern all task decisions, as described below.

Of course, the common factors in the model have to be augmented by factors specific to each different task. Next we describe the model assumptions for each of the three transfer tasks.

Event recognition: During list study incomplete and error prone event traces are stored for each presented character. A test character is encoded by retrieval from knowledge (thereby incorporating both types of frequency dependence) and added to the list and task context cues. The resultant set of features is used to probe event memory. Traces are activated according to similarity to the probe, and include both event traces for the just studied list, and event traces from the training sessions (especially the most recent training session). Each activated trace produces a likelihood ratio, and the average likelihood ratio gives the odds that the test item had been studied. An old decision is given if the odds is greater than 1.0 (or greater than some other criterion that is estimated-- both give a good account of the results). In Experiment 2 it turns out that the frequency effects that are observed at immediate testing (worse performance for high frequency and mirror effects) are due to the activation of traces from the last training session -- more such traces are activated for high frequency probes because more high frequency items were trained in each session. In Experiment 1 this factor was augmented by an additional factor: similarity of probe to list traces varied with frequency because similarity was frequency dependent due to the visual search training. However in Experiment 2 similarity did not vary with frequency due to training with character matching, so frequency effects were due solely to activation of training session traces. This model predicted that frequency effects would drop as delay between training sessions and subsequent study-test increased. This prediction matched the findings.

Pseudo-lexical Decision: The test item produces a gradually increasing number of features extracted from the various knowledge traces as time passes (over several hundred ms). The probe of knowledge gradually becomes richer over time, and the various knowledge traces produce likelihood ratios in accord with their match or similarity to the growing probe. At each moment the average likelihood ratio gives the odds that the item had been studied. When and if the odds exceeds a positive criterion a 'studied' response is given, and when and if the odds drops below a negative criterion, a 'not studied' response is given. The model predicts strong effects of frequency in both experiments and at both immediate and delayed testing, as found.

Forced-choice Perceptual Identification: The briefly presented and masked character that is tested produces a few features extracted from knowledge traces. We approximate this with the presence or absence of a single high level feature (prior research showed low level physical features are ignored in decision making because the masks produce too much noise to make such features useful. The probability of extracting a high level feature is frequency dependent, naturally predicting better performance for higher frequency targets. Why do higher frequency foils produce better performance? We assume sophisticated guessing: When no high level feature has been extracted there is a tendency to guess that a low frequency item had been presented (on the reasonable basis that a high frequency target would more likely have produced a perceived feature, so that the absence of such suggests that a low frequency item had been flashed).

This very abbreviated summary gives the essence of the model and its application. The simplified model as described produced a good qualitative account of the results, and a decent quantitative account as well (although the parameter search was not continued to the point where the best parameter values were located). The in-press article should be read for a much fuller accounting and for extensions to other paradigms and tasks.

A Follow-up Study (not yet submitted):

An important feature of the model is the incorporation in storage of features of nearby events. Does this occur automatically, due to the mere simultaneous presence of such features in short-term memory or is some active process needed to link features together? My graduate student Greg Cox tested this idea with use of a character matching training task, but with two characters presented simultaneously and successively: Two characters appeared and then were repeated shortly thereafter. Two independent decisions were required: Did the left hand character vary in orientation, size, or contrast? Did the right hand character do so? In critical conditions the same two characters co-occurred on many trials. In addition, in case these co-occurring characters would become linked, we built into the design various structures that linked multiple characters (e.g. one simple structure we used is a linear order: A would co-occur with B, B with C, C with D etc.). Various transfer tasks showed that physically co-occurring characters developed some linkage but the effect was not strong and did not extend to indirect linkages. This result demonstrates the critical role of attention and control processes in producing knowledge structures.

Studies of Statistical Learning:

SARKAE is in essence a statistical learning model: Knowledge develops through accumulation of features across different (similar) events. To explore aspects of statistical learning we (faculty colleague Chen Yu and my graduate student George Kachergis) carried out and published several studies of cross modal statistical learning, and produced a model to account for the findings. A typical paradigm used 18 successive study trials. Each trial would have four successive auditory pseudo-words and four simultaneous novel objects that stayed present throughout the trial. There were 'correct' pairings in the sense that such a pairing would always occur together when either member did. However, any one trial does not specify which four of the sixteen possible pairings are correct. The inference about the correct pairings is made statistically across different trials. E.g. if one develops a matrix of co-occurrences across many trials, the pairs that have the highest counts would be the correct ones. In addition there are trial to trial inferences that can be made, especially if the observer assumes that there is a one-to-one mapping of sounds to objects. E.g. If Sound A and Object 1 occur on both trials n and n+1, but all other sounds and objects change across these trials, then one can infer that A-1 is a correct pairing. A variety of othe inferences are also possible. For example, after some pairings have been learned then they can be ignored as possible members of other pairings in an otherwise ambiguous trial. E.g. if one has learned pairings B-2, C-3, and D-4, then a subsequent trial with these six items and two unlearned items allows the inference that the unlearned ones are paired. The various studies published in this series explored these points, and allowed development of a model that showed the way the various inferences combined to produce the observed performance.

Studies of Test Effects:

A critical and key assumption of SARKAE is the storage of all events (in both event traces and knowledge traces). Such storage must include test events. One implication is the beneficial effects of repeated testing of the same item, especially at appropriate selected intervals. This line of research has been explored extensively in recent years by others. A different implication is the harm caused by testing of some items upon retrieval of other items. This line of research has not been explored much and is the basis for several studies by Amy Criss, Ken Malmberg, and students in my lab. The idea is simple: Following study of a list of items for subsequent recognition, targets and foils are tested successively for old/new decisions. The testing can be done singly or as forced choice with each trial containing one target and one foil. The model implies that these test items are stored. A subsequent test of an item (target or foil) not previously tested will tend to activate stored test traces (e.g. produce a significant likelihood ratio), largely because the test context is common to the current memory probe and the stored test traces. Such activation will reduce recognition performance in the same way that other sorts of trace activations do (like study list traces, an pre-experimental traces of the test items). Such findings were obtained and verified in a series of studies that varied stimulus materials, delay of testing, single item and forced choice testing, and changes of categories of test items during testing. A simple version of the SARKAE model was shown to account for the results.

4. Significance of Results and Implications for Science:

The research that was carried out, and in particular the large scale framework we developed for the interrelation of event storage and retrieval and knowledge storage and retrieval represents a major advance in the field. Linking quite different memory and perception tasks, and event memory tasks and knowledge memory tasks, is quite unusual and in itself represents a major advance. Because the Psychological Review article has yet to appear, the impact on science and on research and theory in memory, has yet to take place, but there is good reason to believe the theory we danced will have a major impact on future developments and theory formation in the field.

There were of course many advances, both empirical and theoretical that are important but more narrowly focused. These include the roles of frequency of occurrence in life and training, the role of list traces and pre-list traces in memory, the processes of inference in perceptual recognition, the persistence of knowledge across delays, coupled with the decay of event related effects based on context, the mechanisms of statistical learning, and the effects of storing traces during testing. Each of these should have some important impact on major but specialized domains of research in the field.

5. Publications Resulting from this Research (* partially relevant/supported)

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009) Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.) *Proceedings of the 31st Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009) Temporal contiguity in cross-situational statistical learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.) *Proceedings of the 31st Annual Conference of the Cognitive Science Society Austin, TX: Cognitive Science Society.*

*Sanborn, A. N., Griffiths, T. L., & Shiffrin, R. M. (2010). Uncovering mental representations with Markov chain Monte Carlo. Cognitive Psychology, 60, 63-106.

Kachergis, G., Yu, C., & Shiffrin, R. (2010). Adaptive constraints and inference in cross-situational word learning. *In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the 32nd Annual Conference of the Cognitive Science Society (pp. xxx). Austin, TX: Cognitive Science Society.*

Kachergis, G., Yu, C., & Shiffrin, R. (2010). Cross-situational statistical learning: Implicit or intentional? *In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the 32nd Annual Conference of the Cognitive Science Society*

Cox, G. E., & Shiffrin, R. M. (2011). Criterion setting and the dynamics of recognition memory. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Criss, A. H., Malmberg, K.J., & Shiffrin, R.M. (2011). Output interference in recognition memory testing. *Journal of Memory and Language, 64(4), 316-326.*

*Kachergis, G., Recchia, G., & Shiffrin, R. M. (2011). Adaptive magnitude and valence biases in a dynamic memory task. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Malmberg, K. J., Criss, A. H., Gangwani, T. H., & Shiffrin, R. M. (2012). Overcoming the negative consequences of interference that results from recognition memory testing. *Psychological Science,* **23(2),** 115-119.

Kachergis, G., Yu, C., & Shiffrin, R.M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychonomic Bulletin & Review*, **19(2),** 317-324.

Cox, G. E., & Shiffrin, R. M. (2012). Criterion setting and the dynamics of recognition memory. *Topics in Cognitive Science*, **4(1)**, 135-150.

Kachergis, G., Yu, C., and Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, **5**, 200-213.
(based on the Marr Prize for the best student paper at the 34th annual meeting of the Cognitive Science Society, Sapporo, Japan.)

Kachergis, G., Yu, C., and Shiffrin, R. M. (in press). Cross-situational word earning is better modeled by associations than hypotheses. IEEE Conference on Development and Learning-EpiRob 2012. (awarded 'Best Experiment Combined with Computational Model)

Shiffrin, R. M. (in press). Memory. In Luo, Y, & Lu, Z.-L. (Eds.) *Progress in Cognitive Science: From Cellular Mechanisms to Computational Theories*. Peking University Press.

Nelson, A. B. and Shiffrin, R.M. (in press). The co-evolution of knowledge and event memory. *Psychological Review*.